ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC KHOA HỌC

TRẦN THỊ THU HƯỜNG

PHƯƠNG PHÁP BÌNH PHƯƠNG TỐI THIỂU TOÀN PHẦN

LUẬN VĂN THẠC SĨ TOÁN HỌC

THÁI NGUYÊN - 2019

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC KHOA HỌC

----------------✍📖✍---------------

TRẦN THỊ THU HƯỜNG

PHƯƠNG PHÁP BÌNH PHƯƠNG TỐI THIỂU TOÀN PHẦN

Chuyên ngành: Toán ứng dụng
Mã số : 8 46 01 12

LUẬN VĂN THẠC SĨ TOÁN HỌC

NGƯỜI HƯỚNG DẪN KHOA HỌC
TS. Nguyễn Thị Ngọc Oanh

THÁI NGUYÊN - 2019

Mục lục

Trang

Lời cảm ơn

Luận văn được hoàn thành dưới sự chỉ bảo và hướng dẫn tận tình của TS. Nguyễn Thị Ngọc Oanh. Cô đã dành nhiều thời gian hướng dẫn và giải đáp thắc mắc của tôi trong suốt quá trình làm luận văn. Tôi xin bày tỏ lòng biết ơn sâu sắc đến cô.

Tôi xin gửi tới các thầy, cô trong khoa Toán - Tin của Trường Đại học Khoa học - Đại học Thái Nguyên lời cảm ơn sâu sắc nhất về công lao dạy dỗ trong suốt quá trình học tập tại trường.

Cuối cùng tôi xin cảm ơn Ban giám hiệu trường THPT Quế Võ số 1, tập thể các thầy cô giáo trong tổ Toán-Tin của Trường, gia đình, bạn bè và người thân đã quan tâm, tạo điều kiện, động viên, cổ vũ để tôi có thể hoàn thành nhiệm vụ của mình.

Thái Nguyên, ngày...tháng 4 năm 2019

Học viên

Trần Thị Thu Hường

Lời nói đầu

Phương pháp bình phương tối thiểu toàn phần (Total least square -TLS) được giới thiệu bởi Golub và Van Loan ....như một kỹ thuật giải các hệ "quá xác định" (overdetermined system), tức là những hệ có số phương trình nhiều hơn số ẩn có dạng $AX \approx B$, với các ma trận $A \in \mathbb{R}^{m \times n}$ và $B \in \mathbb{R}^{m \times d}$ cho trước và $X \in \mathbb{R}^{n \times d}$ chưa biết, đang cần tìm. Với $m > n$, nói chung không có nghiệm chính xác X cho xấp xỉ đang tìm. Phương pháp bình phương tối thiểu toàn phần là sự tổng quát quá của phương pháp bình phương tối thiểu (Least square -LS) khi cả ma trận A và B đều bị nhiễu.

Phương pháp bình phương tối thiểu toàn phần là một kỹ thuật hiệu quả để ước lượng tham số tuyến tính. Bài toán ước lượng tham số tuyến tinh đã dẫn tới một lớp rộng rãi các lĩnh vực khoa học như xử lý tín hiệu, điều khiển tự động,..., và nhiều ứng dụng khác trong kỹ thuật, thống kê, vật lý, kinh tế, sinh học,.... Nó được bắt đầu bằng một phương trình tuyến tính sau

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta \tag{0.1}$$

trong đó $\alpha_1, \ldots, \alpha_n$ và β là các biến và $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$ là véc tơ tham số đặc trưng cho hệ (0.1). Bài toán đặt ra là từ dữ kiện đo được nào đó của các biến, ta xác định một ước lượng của tham số chưa biết theo một cách đúng nhất.

Gọi $A \in \mathbb{R}^{m \times n}$ là ma trận có hàng thứ i tương ứng chứa các dữ kiện $\alpha_i, i = 1, \ldots, n$ và véc tơ $b \in \mathbb{R}^m$ chứa β. Khi đó ta có hệ

$$Ax = b \tag{0.2}$$

là hệ gồm m phương trình và n ẩn.

Với cách tiếp cận bài toán bằng phương pháp bình phương tối thiểu thì ma trận A của các dữ kiện đo α_i (vế trái của (0.2)) được giả sử là chính xác, không có sai số. Do đó, các sai số đều được tạo ra bởi véc tơ quan sát b (tức là vế phải của (0.2)). Tuy nhiên điều giả sử này không mang tính thực tế bởi tất các các sai số của mô hình, sai số đo đạc cũng được tạo ra bởi các dữ kiện của ma trận A. Tiếp cận bình phương tối thiểu toàn phần mang ý nghĩa thực tế hơn khi xét sai số được tạo ra cả ở véc tơ quan sát b và ma trận dữ liệu A.

Để minh họa tính hữu hiệu của phương pháp bình phương tối thiểu toàn phần so với phương pháp bình phương tối thiểu, ta có thể xét một ví dụ trong trường hợp ta có một tham số, tức là với $n = 1$. Phương trình (0.1) trở thành

$$\alpha x = \beta. \tag{0.3}$$

Bài toán đặt ra là từ m đo đạc của các biến α và β, ta đi ước lượng tham số x. Ta sẽ đi giải hệ (0.2) với $A = [a_1, \dots, a_m]^T$ và $b = [b_1, \dots, b_m]^T$ trong đó

$$a_i = a_i^0 + \Delta a_i, \tag{0.4}$$

$$b_i = b_i^0 + \Delta b_i,\ i = 1, \dots, m, \tag{0.5}$$

với $\Delta a_i, \Delta b_i$ là các sai số cộng vào các dữ kiện chính xác a_i^0, b_i^0 của các biến α, β.

Nếu α là đo được chính xác, tức là $\Delta a_i = 0$, khi đó sai số chỉ xảy ra khi đo β chứa trong vế phải b. Trong trường hợp này ta có thể sử dụng phương pháp bình phương tối thiểu để giải phương trình (0.2), tức là ta cực tiểu hóa tổng các bình phương có dạng

$$\sum_{i=1}^{m} (b_i - a_i x)^2.$$

Trong trường hợp này, bằng cách khai triển tổng trên thì cực tiểu hay ước lượng x' của x được cho bởi công thức

$$x' = \frac{\sum_{i=1}^{n} a_i b_i}{\sum_{i=1}^{n} a_i^2}$$

Nếu β không có sai số, tức là $\Delta b_i = 0$, ta viết lại phương trình (0.1) dưới dạng

$$\frac{\beta}{x} = \alpha.$$

Sử dụng phương pháp bình phương tối thiểu, bằng cách cực tiểu hóa tổng bình phương sai phân ta thu được ước lượng tốt nhất x'' như sau

$$x' = \frac{\sum_{i=1}^{n} b_i^2}{\sum_{i=1}^{n} a_i b_i}.$$

Tuy nhiên trong thực tế, các biến được đo luôn có sai số, nghĩa là $\Delta a_i \neq 0$ và $\Delta b_i \neq 0$. Trong trường hợp này, ta cần sử dụng phương pháp bình phương tối thiểu toàn phần để nghiên cứu bài toán, cách tiếp cận bài toán bằng phương pháp này mang nhiều ý nghĩa thực tế hơn.

Luận văn được chia làm hai chương. Chương 1 giới thiệu về phương pháp bình phương tối thiểu toàn phần cùng một số kiến thức liên quan như: Phân tích giá trị kì dị, nghiệm bình phương tối thiểu, nghiệm bình phương tối thiểu có chuẩn nhỏ nhất, nghiệm bình phương tối thiểu toàn phần, Định lý về nghiệm bình phương tối thiểu toàn phần,.... Chương 2 của luận văn trình bày một số ứng dụng của phương pháp bình phương tối thiểu toàn phần trong nghiên cứu hồi quy đơn tuyến tính và hồi quy phi tuyến. Một số ví dụ số được trình bày để minh họa cho tính hiệu quả của phương pháp.

Chương 1

Giới thiệu về phương pháp bình phương tối thiểu toàn phần

Trong chương này, chúng tôi trình bày một số kiến thức liên quan tới phương pháp bình phương tối thiểu, phương pháp bình phương tối thiểu toàn phần và có minh họa một vài ví dụ khi sử dụng hai phương pháp nói trên.

1.1. Một số ký hiệu và kiến thức cơ bản

• Ma trận chuyển vị của ma trận A đươc ký hiệu là A^T.

• $R(S)$ (tương ứng $R_r(S)$) là không gian sinh bởi các cột (tương ứng các hàng) của ma trận S, $N(S)$ ký hiệu là không gian véc tơ không hoặc là nhân của S.

• Ma trận đường chéo A cỡ $m \times n$ kí hiệu là

$$A = \mathrm{diag}(\alpha_1, \cdots, \alpha_p), \text{ với } p = \min\{m, n\}$$

với $a_{ij} = 0$ với $i \neq j$ và $a_{ii} = \alpha_i$ với $i = 1, \ldots, p$.

• Ma trận đơn vị cỡ $m \times m$ được kí hiệu là I_m hay đơn giản là I.

• Cho phương trình

$$AX = B \tag{1.1}$$

trong đó A là ma trận cỡ $m \times n$, X là ma trận cỡ $n \times d$, và B là ma trận cỡ $m \times d$.

Nghiệm bình phương tối thiểu có chuẩn nhỏ nhất của phương trình (1.1) được ký hiệu là X', nghiệm bình phương tối thiểu toàn phần có chuẩn nhỏ nhất được ký hiệu là $\hat{X}$.

Trong trường hợp $d = 1$, ta có thể kí hiệu các ma trận X, B bởi các véc tơ x, b tương ứng.

• Chuẩn Frobenius của ma trận M cỡ $m \times n$ được định nghĩa bởi

$$\|M\|_F = \sqrt{\sum_{i=1}^{n}\sum_{i=1}^{m} m_{ij}^2} = \sqrt{\mathrm{tr}(M^T M)}, \tag{1.2}$$

trong đó $\mathrm{tr}(M^T M)$ là vết của ma trận $M^T M$.

• Chuẩn 2 hay chuẩn Euclide của véc tơ y trong không gian n chiều được định nghĩa bởi

$$\|y\|_2 = \sqrt{\sum_{i=1}^{n} y_i^2}. \tag{1.3}$$

• Phân tích giá trị kì dị - SVD.

Ký hiệu SVD của ma trận A cỡ $m \times n,\ m > n$ có dạng

$$A = U'\Sigma' V'^T, \tag{1.4}$$

trong đó

$$U' = [U_1'; U_2'],\ U_1' = [u_1', \cdots, u_n'],\ U_2' = [u_{n+1}', \cdots, u_m'], u_i' \in \mathbb{R}^m,\ U'^T U' = I_m,$$
$$V' = [v_1', \cdots, v_n'],\ v_i' \in \mathbb{R}^n,\ V'^T V' = I_n,$$
$$\Sigma' = \mathrm{diag}(\sigma_1', \cdots, \sigma_n') \in \mathbb{R}^{m\times n},\ \sigma_1' \geq \cdots \geq \sigma_n' \geq 0.$$

Ta cũng kí hiệu SVD của ma trận $[A; B]$ cỡ $m \times (n+d),\ m > n$ có dạng

$$[A; B] = U\Sigma V^T, \tag{1.5}$$

trong đó

$$U = [U_1; U_2],\ U_1 = [u_1, \cdots, u_n],\ U_2 = [u_{n+1}, \cdots, u_m], u_i \in \mathbb{R}^m,\ U^T U = I_m,$$

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} = [v_1, \cdots, v_{n+d}],\ v_i \in \mathbb{R}^{n+d},\ V^T V = I_{n+d},$$

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix} = \mathrm{diag}(\sigma_1, \cdots, \sigma_{n+t}) \in \mathbb{R}^{m\times(n+d)},\ t = \min\{m-n, d\},$$

$$\Sigma_1 = \mathrm{diag}(\sigma_1, \cdots, \sigma_n) \in \mathbb{R}^{n\times n},$$

$$\Sigma_1 = \mathrm{diag}(\sigma_{n+1}, \cdots, \sigma_{n+t}) \in \mathbb{R}^{(m-n)\times d},$$

và $\sigma_1 \geq \cdots \geq \sigma_{n+t} \geq 0$.

1.2. Phương pháp bình phương tối thiểu

Phương pháp bình phương tối thiểu (LS) là một cách tiếp cận phổ biến khi giải hệ tuyến tính quá xác định, tức là số phương trình nhiều hơn số ẩn. Nói chung các hệ như vậy không có nghiệm nhưng ta tìm nghiệm xấp xỉ của hệ bằng cách cực tiểu hóa tổng bình phương của sai số được tạo thành khi giải mỗi phương trình đơn lẻ.

Xét bài toán tìm $x \in \mathbb{R}^n$ thỏa mãn phương trình

$$Ax = b,$$

trong đó $b \in \mathbb{R}^m$ đã biết và dữ liệu $A \in \mathbb{R}^{m\times n}$. Khi số phương trình nhiều hơn số ẩn, tức là $m > n$ thì hệ được gọi là *hệ quá xác định*. Trong trường hợp $b \notin R(A)$ thì hệ quá xác định không có nghiệm chính xác, do vậy ta sử dụng kí hiệu $Ax \approx b$.

Nghiệm bình phương tối thiểu được trình bày trong định nghĩa dưới đây.

Định nghĩa 1.1 Xét bài toán tìm cực tiểu

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|_2, \quad A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m. \tag{1.6}$$

Cực tiểu x' bất kỳ được gọi là nghiệm bình phương tối thiểu của hệ $Ax \approx b$.